(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 August 2001 (30.08.2001)

PCT

(10) International Publication Number
WO 01/62802 A1

(51) International Patent Classification[7]: C07K 16/40, C12N 9/64, A61K 39/00

(21) International Application Number: PCT/GB01/00819

(22) International Filing Date: 26 February 2001 (26.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0004533.6 25 February 2000 (25.02.2000) GB

(71) Applicant *(for all designated States except US)*: UNIVERSITY OF NOTTINGHAM [GB/GB]; University Park, Nottingham NG7 2RD (GB).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: PRITCHARD, David, Idris [GB/GB]; University of Nottingham, University Park, Nottingham NG7 2RD (GB).

(74) Agent: GILL JENNINGS & EVERY; Broadgate House, 7 Eldon Street, London EC2M 7LH (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/62802 A1

(54) Title: TREATMENT OF HOOKWORM INFECTION

(57) Abstract: Vaccine compositions for the treatment of hookworm infections comprise antigenic fragments of aspartyl proteinases obtainable from *Necator americanus*.

TREATMENT OF HOOKWORM INFECTION

Field of the Invention

This invention relates to the production of vaccine compositions to treat parasitic infection, in particular to treat infection of the hookworm *Necator americanus*.

Background to the Invention

The human hookworm *Necator americanus* is a human pathogen that invades the body by penetrating the skin, and causes debilitating iron deficiency anaemia at low infection intensity.

Treating infection using pharmaceuticals can be carried out but the effect is often transient, and the treatment is costly. Hookworm vaccines have been used successfully to control the pathology associated with canine infections. However, protection in this case was induced by exposure to live γ-radiation-attenuated infective larvae, and this treatment is unlikely to be acceptable for human use.

Matthews, Z Parasitenkd, 1982; 68: 81-86, discloses that cellular destruction of the skin during larval penetration through the epidermis is effected by an undefined enzymatic process. It was shown subsequently that serine, and possibly cysteinyl, proteinases were responsible for skin penetration. However, the precise role for each of these proteinases was not defined.

Brown *et al*, Am. J. Trop. Med. Hyg., 1999; 60(5): 840-847, identifies aspartyl proteinase activity to be important for larval stage skin penetration. No specific aspartyl proteinase is identified. Treatment to prevent skin penetration is proposed using general aspartyl proteinase inhibitors.

Summary of the Invention

The present invention is based on the realisation that the aspartyl proteinase of *Necator americanus* is a viable target for vaccine therapy.

Not only may aspartyl proteinases be important for larval stage hookworm skin penetration, but it is now appreciated that they may be important in the maintenance of the mature parasite life-cycle. This is based on the finding that adult parasites appear to depend predominantly on aspartyl proteinase activity to digest host haemoglobin and fibrinogen, which may be important to the maintenance of a haematophagous existence in the gut.

According to a first aspect of the invention, a vaccine composition comprises an aspartyl proteinase obtainable from the hookworm *Necator americanus*, or an antigenic fragment thereof.

According to a second aspect of the invention, a vaccine composition comprises a polynucleotide that encodes an aspartyl proteinase obtainable from the hookworm *Necator americanus*, or an antigenic fragment thereof.

According to a third aspect of the invention, an antibody is raised against an aspartyl proteinase, as defined above, and may be used in therapy or diagnosis.

According to a fourth aspect of the invention, an aspartyl proteinase comprises the amino acid sequence identified herein as SEQ ID NO. 6. This aspartyl proteinase is found only in the adult hookworm *Necator americanus*, and is structurally different from that found in larval forms.

In contrast to the prior art, the present invention provides means to treat an existing infection or to prevent infection. It was not at all apparent, until now, that an effective vaccine could be produced using an aspartyl proteinase as the antigenic fragment, and that this would be effective against hookworm infection. The identification of structurally different aspartyl proteinases is an important aspect in the development of the vaccine compositions, as is the finding that these are structurally different from the human aspartyl proteinases. It is particularly surprising that the adult hookworm

contains structurally different aspartyl proteinases to that of the larval hookworm.

Description of the Invention

The present invention provides treatments for parasitic infection, in particular from infection by human hookworm, e.g. *Necator americanus*. However, it is not intended to restrict the treatments to infections of a human host, and the present invention extends to veterinary treatment of animal infections, for example, by the related canine hookworm *Ancylostoma caninum*, or the sheep hookworm *Haemonchus contortus*.

Specific aspartyl proteinases are identified herein on the basis of polynucleotide and amino acid sequences (identified herein as SEQ ID NOS. 2, 3, 4 and 6). Homologues to these sequences, with at least 60%, preferably at least 80% or 90%, sequence identity or similarity (measured across the complete sequence) are also within the scope of the invention.

The terms "similarity" and "identity" are known in the art. The use of the term "identity" refers to a sequence comparison based on identical matches between correspondingly identical positions in the sequences being compared. The term "similarity" refers to a comparison between amino acid sequences, and takes into account not only identical amino acids in corresponding positions, but also functionally similar amino acids in corresponding positions. Thus, similarity between polypeptide sequences indicates functional similarity, in addition to sequence similarity.

Levels of identity between gene sequences and levels of identity or similarity between amino acid sequences can be calculated using known methods. In relation to the present invention, publicly available computer-based methods for determining identity and similarity include the BLASTP, BLASTN and FASTA (Atschul *et al.*, J. Molec. Biol., 1990; 215:403-410), the BLASTX program available from NCBI, and the Gap program from Genetics Computer Group, Madison

WI. The levels of similarity and identity referred to herein, are calculated with reference to the Gap program, with a Gap penalty of 12 and a Gap length penalty of 4 for determining the amino acid sequence comparisons, and a Gap penalty of 50 and a Gap length penalty of 3 for the polynucleotide sequence comparisons.

The aspartyl proteinases according to the invention may be purified and isolated by methods known in the art. In particular, having identified the gene sequence or the N-terminal sequence, it will be possible to use recombinant techniques to express the genes in a suitable host.

Active fragments of the proteins and polynucleotides are those that retain the biological function of the protein or polynucleotide. For example, when used as part of the vaccine to elicit an immune response, the fragment will be of sufficient size, such that antibodies generated in response to the fragment will be specific for that aspartyl proteinase and will not, for example, cross-react with the natural aspartyl proteinases of the patient. Typically, the fragment will be at least 30 nucleotides (10 amino acids) in size, preferably 60 nucleotides (20 amino acids) and most preferably greater than 90 nucleotides (30 amino acids) in size.

It should also be understood that the invention encompasses modifications made to the proteins and polynucleotides identified herein which do not significantly alter the biological function. It will be apparent to the skilled person that the degeneracy of the genetic code can result in polynucleotides with minor base changes from those specified herein, but which nevertheless encode the same proteins. Complementary polynucleotides are also within the invention. Conservative replacements at the amino acid level are also envisaged, i.e. different acidic or basic amino acids may be substituted without substantial loss of function.

The preparation of vaccines based on the aspartyl proteinases will be apparent to those skilled in the art.

Vaccine compositions can be formulated with suitable carriers or adjuvants, e.g. alum, as necessary or desired, to provide effective immunisation against infection.

It is preferred that the vaccines are prepared in order to elicit a T-helper type-2 cell response. The adjuvant may therefore comprise components that influence this, and it may be preferable not to include adjuvants comprising bacterial components which induce T-helper type-1 cell responses.

More generally, and as is well known to those skilled in the art, a suitable amount of an active component of the invention can be selected, for therapeutic use, as can suitable carriers or excipients, and routes of administration. These factors would be chosen or determined according to known criteria such as the nature/severity of the condition to be treated, the type and/or health of the subject etc.

The vaccine may comprise an antigenic fragment of an aspartyl proteinase characterised as present in the larval stage, or alternatively, present in the adult stage. In a preferred embodiment, the vaccine composition comprises a combination of an antigenic fragment derived from a larval stage aspartyl proteinase and an antigenic fragment derived from an adult stage aspartyl proteinase. This offers maximum protection as it targets separate stages of hookworm infection.

In a further preferred embodiment, the aspartyl proteinase from which the vaccine may be prepared, is encoded by the DNA sequence defined as SEQ ID NO. 1, or SEQ ID NO. 5, or a homologue thereof with at least 60% sequence identity, preferably 80%, and most preferably 95% sequence identity.

The vaccine may also be derived from an aspartyl proteinase characterised as comprising an amino acid sequence shown as SEQ ID NO. 3 or SEQ ID NO. 4.

The vaccine may comprise alternatively a genetic construct that encodes an aspartyl proteinase, or a

fragment thereof. In this embodiment, it may be necessary to prepare the construct to include appropriate regulatory factors, e.g. promoters, in addition to the polynucleotide that encodes the proteinase. Suitable components, including suitable vectors, will be apparent to the skilled person.

The invention will now be further described by way of example only with reference to aspartyl proteinases isolated from *N. americanus*.

Example

Preparation of *N. americanus* larval secretions

Infective larvae were cultured from faecal material as described by Kumar and Pritchard, Int. J. Parasitol, 1992; 22:563-572. Briefly, faecal material obtained from hamsters infected with *N. americanus* was mixed with activated charcoal, 1% (w/v) amphotericin B and water to form a smooth paste which was applied to the upper half of a 5 x 30 cm strip of filter paper. These strips were then suspended in a large glass chromatography tank containing approximately 750 ml of distilled water. The tanks were sealed and incubated at 28°C for 10 days, after which the filter paper strips were carefully removed and discarded. The water containing the larvae was transferred to a measuring cylinder and the larvae allowed to sediment for two hours. After this period the water was aspirated off and the larvae washed twice to remove any faecal contamination. Finally, washed larvae were re-suspended in approximately 20 ml of storage buffer (50 mM Na_2HPO_4, 70 mM NaCl, 15 mM KH_2PO_4, pH7.4). Larvae were stored in the dark at room temperature until required, or for a maximum period of one month.

Excretory-secretory (ES) products were collected as described by Kumar and Pritchard (1992), *supra*. Freshly collected, ensheaved larvae were re-suspended in larval storage buffer and exsheaved by bubbling carbon dioxide through the suspension for two hours at room temperature. Exsheaved larvae were allowed to settle and then washed

extensively with RPMI 1640 containing 100 i.u./ml penicillin, 100 μg/ml streptomycin and 1% amphotericin B under sterile conditions. Following this sterilisation period the larvae were cultured in RPMI 1640 containing the above additives for 72 hours at 37°C, changing the culture medium every 24 hours. ES products collected over the 72 hour period were pooled, dialysed against distilled water, lyophilized and stored at -20°C until required.

Enzyme Purification

Substrate SDS-PAGE was carried out using a method modified from Pritchard *et al*, Parasitology Today, 1990; 6: 154-156. 12% (w/v) SDS-PAGE gels were prepared with the inclusion of 0.1% (w/v) haemoglobin in the resolving gel. 10 μg of the ES products was mixed with an equal volume of non-reducing sample buffer (0.5M Tris, pH 6.8, 5% SDS (w/v), 20% glycerol (w/v), 0.01% bromophenol) and incubated under 37°C for 30 minutes. The sample was then applied to the gel which was then electrophoresed at a constant current of 20 mA. Following electrophoresis, the gels were washed in 2.5% Triton X-100 for one hour at room temperature to renature the enzymes. The gels were then washed in water for 30 minutes, cut into individual strips and incubated for 48 hours at 37°C in 0.1 M sodium phosphate buffer pH 6.5. Proteinase activity was detected by staining gels with Coomassie brilliant blue R250.

The gels revealed three proteinase products at 31kD, 33kD and 35kD.

Larval aspartyl proteinase was purified from the ES products using pepsatin A agarose (Sigma). A 5 ml pepsatin A agarose column was equilibrated with 50 mM sodium acetate pH 5.5. The ES products in 50 mM sodium acetate pH 5.5 were applied to the column at a flow rate of 0.2 ml/min. The column was washed sequentially with 10 ml, 50 mM sodium acetate pH 5.5 followed by 10 ml, 50 mM sodium acetate, 0.5 M sodium chloride pH 5.5. Bound protein was eluted from the column with 15 ml, 500 μM pepsatin A dissolved in 50 mM sodium acetate pH 5.5. One ml fractions were collected and

analysed for protein content and proteolytic activity using FITC-labelled casein. Fractions eluted from the column containing 500 μM pepsatin A were dialysed against distilled water prior to analysis for proteolytic activity.

The aspartyl proteinases present in the purified fractions were sequenced to obtain information on their amino acid and nucleic acid structure. The DNA sequence for one of the larval aspartyl proteinases is shown as SEQ ID NO. 1 and the amino acid sequence is shown as SEQ ID NO. 2. N-terminal sequencing was carried out for two other larval aspartyl proteinases, and the sequences are shown as SEQ ID NOS. 3 and 4.

The measurement of proteinase activity, using FITC-casein as the substrate, revealed that the activity was optimal at pH 6.5. At pH 6.5, proteinase activity was also shown to be inhibited by pepsatin A.

An aspartyl proteinase was also purified from the adult hookworm using techniques similar to those described. This proteinase had an amino acid and nucleic acid sequence significantly different to those obtained from the larval hookworm. The nucleic acid sequence is shown as SEQ ID NO. 5 and the amino acid sequence is shown as SEQ ID NO. 6.

The aspartyl proteinase obtained from the adult form was tested in assays to determine its substrate specificity. It was found that the proteinase cleaved the synthetic peptide substrate ALERTFLSFPT (SEQ ID NO. 7). This synthetic substrate mimics the site at which initial cleavage of haemaglobin by *P. falciparum* aspartic proteinases is known to occur. Adult aspartyl proteinase may therefore be important in the digestion of host haemoglobin and fibrinogen and may therefore be an important factor in anti-coagulation, maintaining the hookworm in the host.

CLAIMS

1. A vaccine composition comprising an aspartyl proteinase obtainable from the hookworm *Necator americanus* or an antigenic fragment thereof.

2. A composition according to claim 1, wherein the proteinase or fragment comprises part or all of any of the amino acid sequences defined herein as SEQ ID NOS. 2, 3, 4 and 6, or a homologue thereof with at least 60% sequence similarity.

3. A composition according to claim 1 or claim 2, wherein the fragment is at least 30 amino acids.

4. A composition according to any preceding claim, comprising both adult and larval aspartyl proteinases or antigenic fragments thereof.

5. A vaccine composition comprising a polynucleotide that encodes an aspartyl proteinase obtainable from the hookworm *Necator americanus* or an antigenic fragment thereof.

6. A composition according to claim 5, wherein the polynucleotide comprises SEQ ID NO. 1 or SEQ ID NO. 5, or a homologue thereof with at least 60% sequence identity.

7. A composition according to claim 5 or claim 6, comprising polynucleotides encoding each of adult and larval aspartyl proteinases or antigenic fragments thereof.

8. Use of an aspartyl proteinase as defined in any of claims 1 to 4, or a polynucleotide as defined in any of claims 5 to 7, in the manufacture of a vaccine composition for the treatment of a hookworm infection.

9. Use according to claim 8, wherein the infection is a *Necator americanus* infection.

10. Use according to claim 8, wherein the infection is an *Ancylostoma caninum* infection.

11. Use according to claims 8, wherein the infection is an *Haemonchus contortus* infection.

12. An antibody raised against an aspartyl proteinase as defined in any of claims 1 to 3.

13. An aspartyl proteinase obtainable from *Necator americanus*, encoded by a gene comprising the polynucleotide identified herein as SEQ ID NO. 5.

14. An aspartyl proteinase according to claim 13, for therapeutic use.

SEQUENCE LISTING

<110> University of Nottingham

<120> Treatment of Hookworm Infection

<130> REP06004WO

<140> (not yet known)
<141> 2001-02-26

<150> 0004533.6
<151> 2000-02-25

<160> 7

<170> PatentIn Ver. 2.1

<210> 1
<211> 1341
<212> DNA
<213> Necator americanus

<220>
<221> CDS
<222> (1)..(1341)

<400> 1

```
atg gct cga ctt gta ttc cta ctc gta cta tgt act ctg gct gca gca   48
Met Ala Arg Leu Val Phe Leu Leu Val Leu Cys Thr Leu Ala Ala Ala
  1               5                  10                  15

agc gtt cat cga cga ctc ttt cat caa gct cgt cgt cat gtg aca tcg   96
Ser Val His Arg Arg Leu Phe His Gln Ala Arg Arg His Val Thr Ser
             20                  25                  30

gta tcg ctt tcg cgt cag cca aca ctt cgt gaa cga ctg atc gca agt   144
Val Ser Leu Ser Arg Gln Pro Thr Leu Arg Glu Arg Leu Ile Ala Ser
         35                  40                  45

ggc agt tgg gag gat tac cag aaa caa cgc tac cat tat cga aag aaa   192
Gly Ser Trp Glu Asp Tyr Gln Lys Gln Arg Tyr His Tyr Arg Lys Lys
     50                  55                  60

att cta gca aaa tat gct gct aac aaa gcg tca aag tta caa tct gca   240
Ile Leu Ala Lys Tyr Ala Ala Asn Lys Ala Ser Lys Leu Gln Ser Ala
 65                  70                  75                  80
```

```
aac gag atc gat gaa ttg ctc cgg aac tat atg gat gca caa tac tat   288
Asn Glu Ile Asp Glu Leu Leu Arg Asn Tyr Met Asp Ala Gln Tyr Tyr
                 85                  90                  95

ggt gtc atc caa att ggg act cca gct cag aat ttc act gtg atc ttc   336
Gly Val Ile Gln Ile Gly Thr Pro Ala Gln Asn Phe Thr Val Ile Phe
            100                 105                 110

gac acg ggt tcc tca aat cta tgg gta ccg tca aga aag tgt cca ttc   384
Asp Thr Gly Ser Ser Asn Leu Trp Val Pro Ser Arg Lys Cys Pro Phe
        115                 120                 125

tat gac att gca tgt atg ctt cat cat cgt tat gac tcc gga gcc tcg   432
Tyr Asp Ile Ala Cys Met Leu His His Arg Tyr Asp Ser Gly Ala Ser
    130                 135                 140

tca acc tac aag gaa gat ggg cgc aag atg gct att cag tat gga act   480
Ser Thr Tyr Lys Glu Asp Gly Arg Lys Met Ala Ile Gln Tyr Gly Thr
145                 150                 155                 160

gga tct atg aaa gga ttc att tct aag gat att gtt tgt att gct gga   528
Gly Ser Met Lys Gly Phe Ile Ser Lys Asp Ile Val Cys Ile Ala Gly
                165                 170                 175

att tgc gct gaa gaa caa cct ttc gcg gag gct aca agt gaa cct ggt   576
Ile Cys Ala Glu Glu Gln Pro Phe Ala Glu Ala Thr Ser Glu Pro Gly
            180                 185                 190

ctt aca ttt atc gct gct aag ttt gat gga atc ctt gga atg gca ttc   624
Leu Thr Phe Ile Ala Ala Lys Phe Asp Gly Ile Leu Gly Met Ala Phe
        195                 200                 205

ccg gaa att gct gtt ctc ggt gta act cct gtc ttc cat acg ttc att   672
Pro Glu Ile Ala Val Leu Gly Val Thr Pro Val Phe His Thr Phe Ile
    210                 215                 220

gaa cag aag aaa gtt cct agc cct gtg ttt gct ttc tgg ctg aat agg   720
Glu Gln Lys Lys Val Pro Ser Pro Val Phe Ala Phe Trp Leu Asn Arg
225                 230                 235                 240

aat cca gag tcg gaa att gga gga gag att acc ttt ggt ggt gtg gat   768
Asn Pro Glu Ser Glu Ile Gly Gly Glu Ile Thr Phe Gly Gly Val Asp
                245                 250                 255

acc cga cgt tat gtt gaa cca att aca tgg aca cca gtg aca cgt cgt   816
Thr Arg Arg Tyr Val Glu Pro Ile Thr Trp Thr Pro Val Thr Arg Arg
            260                 265                 270
```

```
gga tat tgg caa ttc aaa atg gat atg gta caa ggt ggt tca tcg tcc   864
Gly Tyr Trp Gln Phe Lys Met Asp Met Val Gln Gly Gly Ser Ser Ser
        275                 280                 285

att gcg tgt ccg aat gga tgc caa gct atc gct gat act ggc act tct   912
Ile Ala Cys Pro Asn Gly Cys Gln Ala Ile Ala Asp Thr Gly Thr Ser
    290                 295                 300

ctt att gct gga ccg aag gca cag gtt gag gca atc cag aaa tat atc   960
Leu Ile Ala Gly Pro Lys Ala Gln Val Glu Ala Ile Gln Lys Tyr Ile
305                 310                 315                 320

gga gca gag ccg ctt atg aaa gga gaa tac atg att cct tgc gac aaa   1008
Gly Ala Glu Pro Leu Met Lys Gly Glu Tyr Met Ile Pro Cys Asp Lys
                325                 330                 335

gta cca tcc ctt cct gat gtt tcg ttc atc atc gat ggc aag acg ttt   1056
Val Pro Ser Leu Pro Asp Val Ser Phe Ile Ile Asp Gly Lys Thr Phe
            340                 345                 350

aca ctc aaa ggg gaa gat tac gtt cta acc gtg aaa gcc gct ggt aaa   1104
Thr Leu Lys Gly Glu Asp Tyr Val Leu Thr Val Lys Ala Ala Gly Lys
        355                 360                 365

tca atc tgt ttg tct ggc ttc atg gga atg gac ttc cca gag aag atc   1152
Ser Ile Cys Leu Ser Gly Phe Met Gly Met Asp Phe Pro Glu Lys Ile
    370                 375                 380

ggc gaa ttg tgg atc ctt gga gat gtt ttc att gga aaa tac tac acc   1200
Gly Glu Leu Trp Ile Leu Gly Asp Val Phe Ile Gly Lys Tyr Tyr Thr
385                 390                 395                 400

gtc ttc gat gtt ggt cag gca cgt gtt gga ttt gct caa gca aag tca   1248
Val Phe Asp Val Gly Gln Ala Arg Val Gly Phe Ala Gln Ala Lys Ser
                405                 410                 415

gaa gat gga ttc cct gtt ggc acc ccc gtt cga aca ttc aga cag ctt   1296
Glu Asp Gly Phe Pro Val Gly Thr Pro Val Arg Thr Phe Arg Gln Leu
            420                 425                 430

cag gaa gac agc gat agc gac gag gac gat gta ttt act ttt taa       1341
Gln Glu Asp Ser Asp Ser Asp Glu Asp Asp Val Phe Thr Phe
        435                 440                 445
```

<210> 2
<211> 446
<212> PRT

<213> Necator americanus

<400> 2

```
Met Ala Arg Leu Val Phe Leu Leu Val Leu Cys Thr Leu Ala Ala Ala
  1               5                  10                  15

Ser Val His Arg Arg Leu Phe His Gln Ala Arg Arg His Val Thr Ser
             20                  25                  30

Val Ser Leu Ser Arg Gln Pro Thr Leu Arg Glu Arg Leu Ile Ala Ser
         35                  40                  45

Gly Ser Trp Glu Asp Tyr Gln Lys Gln Arg Tyr His Tyr Arg Lys Lys
     50                  55                  60

Ile Leu Ala Lys Tyr Ala Ala Asn Lys Ala Ser Lys Leu Gln Ser Ala
 65                  70                  75                  80

Asn Glu Ile Asp Glu Leu Leu Arg Asn Tyr Met Asp Ala Gln Tyr Tyr
                 85                  90                  95

Gly Val Ile Gln Ile Gly Thr Pro Ala Gln Asn Phe Thr Val Ile Phe
            100                 105                 110

Asp Thr Gly Ser Ser Asn Leu Trp Val Pro Ser Arg Lys Cys Pro Phe
        115                 120                 125

Tyr Asp Ile Ala Cys Met Leu His His Arg Tyr Asp Ser Gly Ala Ser
    130                 135                 140

Ser Thr Tyr Lys Glu Asp Gly Arg Lys Met Ala Ile Gln Tyr Gly Thr
145                 150                 155                 160

Gly Ser Met Lys Gly Phe Ile Ser Lys Asp Ile Val Cys Ile Ala Gly
                165                 170                 175

Ile Cys Ala Glu Glu Gln Pro Phe Ala Glu Ala Thr Ser Glu Pro Gly
            180                 185                 190

Leu Thr Phe Ile Ala Ala Lys Phe Asp Gly Ile Leu Gly Met Ala Phe
        195                 200                 205

Pro Glu Ile Ala Val Leu Gly Val Thr Pro Val Phe His Thr Phe Ile
    210                 215                 220

Glu Gln Lys Lys Val Pro Ser Pro Val Phe Ala Phe Trp Leu Asn Arg
225                 230                 235                 240
```

```
Asn Pro Glu Ser Glu Ile Gly Gly Glu Ile Thr Phe Gly Gly Val Asp
                245                 250                 255

Thr Arg Arg Tyr Val Glu Pro Ile Thr Trp Thr Pro Val Thr Arg Arg
            260                 265                 270

Gly Tyr Trp Gln Phe Lys Met Asp Met Val Gln Gly Gly Ser Ser Ser
        275                 280                 285

Ile Ala Cys Pro Asn Gly Cys Gln Ala Ile Ala Asp Thr Gly Thr Ser
    290                 295                 300

Leu Ile Ala Gly Pro Lys Ala Gln Val Glu Ala Ile Gln Lys Tyr Ile
305                 310                 315                 320

Gly Ala Glu Pro Leu Met Lys Gly Glu Tyr Met Ile Pro Cys Asp Lys
                325                 330                 335

Val Pro Ser Leu Pro Asp Val Ser Phe Ile Ile Asp Gly Lys Thr Phe
            340                 345                 350

Thr Leu Lys Gly Glu Asp Tyr Val Leu Thr Val Lys Ala Ala Gly Lys
        355                 360                 365

Ser Ile Cys Leu Ser Gly Phe Met Gly Met Asp Phe Pro Glu Lys Ile
    370                 375                 380

Gly Glu Leu Trp Ile Leu Gly Asp Val Phe Ile Gly Lys Tyr Tyr Thr
385                 390                 395                 400

Val Phe Asp Val Gly Gln Ala Arg Val Gly Phe Ala Gln Ala Lys Ser
                405                 410                 415

Glu Asp Gly Phe Pro Val Gly Thr Pro Val Arg Thr Phe Arg Gln Leu
            420                 425                 430

Gln Glu Asp Ser Asp Ser Asp Glu Asp Asp Val Phe Thr Phe
        435                 440                 445
```

<210> 3
<211> 10
<212> PRT
<213> Necator americanus

<400> 3
Asp Val Ile Pro Gln Val Ala His Asp Tyr

1 5 10

<210> 4
<211> 15
<212> PRT
<213> Necator americanus

<400> 4
Gly Asn Val Val Pro Gln Ala Val Asn Asp Phe Thr Asp Val Gln
1 5 10 15

<210> 5
<211> 1278
<212> DNA
<213> Necator americanus

<220>
<221> CDS
<222> (1)..(1278)

<400> 5
atg cgt tcg ata ctc gtg ttg gtg gct ctg atc gga tgc att gct gcg 48
Met Arg Ser Ile Leu Val Leu Val Ala Leu Ile Gly Cys Ile Ala Ala
1 5 10 15

ggt gta tat aaa atc cca ttg aaa aga atc act ccg ccg atg ata aaa 96
Gly Val Tyr Lys Ile Pro Leu Lys Arg Ile Thr Pro Pro Met Ile Lys
20 25 30

atg ttg aga gct ggt act tgg gaa acg tac gta gaa gga atg agg aag 144
Met Leu Arg Ala Gly Thr Trp Glu Thr Tyr Val Glu Gly Met Arg Lys
35 40 45

aga caa tta cag tta ctg aag gag cac aag gtt cat atc caa gat gta 192
Arg Gln Leu Gln Leu Leu Lys Glu His Lys Val His Ile Gln Asp Val
50 55 60

ctc ggc tat gct aac atg gag tac ctc ggc gaa att act att gga act 240
Leu Gly Tyr Ala Asn Met Glu Tyr Leu Gly Glu Ile Thr Ile Gly Thr
65 70 75 80

cct caa cag aag ttt ctg gtg gtt ttg gac act ggc tcc tcg aat ctg 288
Pro Gln Gln Lys Phe Leu Val Val Leu Asp Thr Gly Ser Ser Asn Leu
85 90 95

```
tgg gtc cct gat gat tca tgc tac aag gag aag aga cct gat aga tgt   336
Trp Val Pro Asp Asp Ser Cys Tyr Lys Glu Lys Arg Pro Asp Arg Cys
            100                 105                 110

cta gta tca aac tgt gat gct gga ctg gtt tgt caa gtc ttc tgt cca   384
Leu Val Ser Asn Cys Asp Ala Gly Leu Val Cys Gln Val Phe Cys Pro
        115                 120                 125

gat cct aaa tgc tgt gaa cat acg aga gaa ttc aag caa gta aac gca   432
Asp Pro Lys Cys Cys Glu His Thr Arg Glu Phe Lys Gln Val Asn Ala
    130                 135                 140

tgc aaa gat aag cat cga ttt gat caa aag aat tcc aac act tat gtt   480
Cys Lys Asp Lys His Arg Phe Asp Gln Lys Asn Ser Asn Thr Tyr Val
145                 150                 155                 160

aaa aca aac aaa aca tgg gca ata gcg tat gga act gga gat gcg agg   528
Lys Thr Asn Lys Thr Trp Ala Ile Ala Tyr Gly Thr Gly Asp Ala Arg
                165                 170                 175

gga ttt ttt gga aga gat aca gtc cgt ttg ggt gct gaa gga aag gat   576
Gly Phe Phe Gly Arg Asp Thr Val Arg Leu Gly Ala Glu Gly Lys Asp
            180                 185                 190

cag ctc gtt att aat gat acg tgg ttc gga caa gca gag cat ata gct   624
Gln Leu Val Ile Asn Asp Thr Trp Phe Gly Gln Ala Glu His Ile Ala
        195                 200                 205

gaa ttt ttc agt aat act ttc ctt gat ggc att ctc gga ctc gct ttt   672
Glu Phe Phe Ser Asn Thr Phe Leu Asp Gly Ile Leu Gly Leu Ala Phe
    210                 215                 220

caa gaa ctg tca gaa gga ggc gtc gct cct cca ata att cgt gcc att   720
Gln Glu Leu Ser Glu Gly Gly Val Ala Pro Pro Ile Ile Arg Ala Ile
225                 230                 235                 240

gac ctt gga ctt ctc gat caa cca ata ttt act gtc tat ttc gaa aat   768
Asp Leu Gly Leu Leu Asp Gln Pro Ile Phe Thr Val Tyr Phe Glu Asn
                245                 250                 255

gtc gga gac aaa gaa ggt gtt tat gga ggt gtt ttc acc tgg ggt ggt   816
Val Gly Asp Lys Glu Gly Val Tyr Gly Gly Val Phe Thr Trp Gly Gly
            260                 265                 270

ctc gat ccc gat cat tgc gaa gat gag gtc aca tat gaa cag cta acc   864
Leu Asp Pro Asp His Cys Glu Asp Glu Val Thr Tyr Glu Gln Leu Thr
        275                 280                 285
```

```
gaa gca act tac tgg cag ttt aga ctt aaa gga gtg tcg tct aag aac   912
Glu Ala Thr Tyr Trp Gln Phe Arg Leu Lys Gly Val Ser Ser Lys Asn
    290                 295                 300

ttc tcg tcg acg gct ggt tgg gaa gca ata tcc gac act ggt acc tcg   960
Phe Ser Ser Thr Ala Gly Trp Glu Ala Ile Ser Asp Thr Gly Thr Ser
305                 310                 315                 320

tta aat gga gcc cct agg ggg ata cta aga agt att gca aga cag tat   1008
Leu Asn Gly Ala Pro Arg Gly Ile Leu Arg Ser Ile Ala Arg Gln Tyr
                325                 330                 335

aat gga cag tac gtc gca tct caa ggt ctc tac gtc gtc gac tgc agt   1056
Asn Gly Gln Tyr Val Ala Ser Gln Gly Leu Tyr Val Val Asp Cys Ser
            340                 345                 350

aaa aat gtg acc gtt gac gtg acc att ggc gac aga aac tac act atg   1104
Lys Asn Val Thr Val Asp Val Thr Ile Gly Asp Arg Asn Tyr Thr Met
        355                 360                 365

act gcg aaa aat ctc gta ctt gaa ata cag gct gat ata tgt att atg   1152
Thr Ala Lys Asn Leu Val Leu Glu Ile Gln Ala Asp Ile Cys Ile Met
    370                 375                 380

gca ttt ttc gaa atg gac atg ttc att gga cca gca tgg att ctt ggc   1200
Ala Phe Phe Glu Met Asp Met Phe Ile Gly Pro Ala Trp Ile Leu Gly
385                 390                 395                 400

gat cca ttt att cga gaa tat tgc aat att cat gac att gaa aag aag   1248
Asp Pro Phe Ile Arg Glu Tyr Cys Asn Ile His Asp Ile Glu Lys Lys
                405                 410                 415

cgg att ggt ttt gca gct gta aaa cat tga                          1278
Arg Ile Gly Phe Ala Ala Val Lys His
            420                 425
```

<210> 6
<211> 425
<212> PRT
<213> Necator americanus

<400> 6

```
Met Arg Ser Ile Leu Val Leu Val Ala Leu Ile Gly Cys Ile Ala Ala
  1               5                  10                  15

Gly Val Tyr Lys Ile Pro Leu Lys Arg Ile Thr Pro Pro Met Ile Lys
```

```
            20                  25                  30

Met Leu Arg Ala Gly Thr Trp Glu Thr Tyr Val Glu Gly Met Arg Lys
         35                  40                  45

Arg Gln Leu Gln Leu Leu Lys Glu His Lys Val His Ile Gln Asp Val
     50                  55                  60

Leu Gly Tyr Ala Asn Met Glu Tyr Leu Gly Glu Ile Thr Ile Gly Thr
 65                  70                  75                  80

Pro Gln Gln Lys Phe Leu Val Val Leu Asp Thr Gly Ser Ser Asn Leu
                 85                  90                  95

Trp Val Pro Asp Asp Ser Cys Tyr Lys Glu Lys Arg Pro Asp Arg Cys
            100                 105                 110

Leu Val Ser Asn Cys Asp Ala Gly Leu Val Cys Gln Val Phe Cys Pro
        115                 120                 125

Asp Pro Lys Cys Cys Glu His Thr Arg Glu Phe Lys Gln Val Asn Ala
    130                 135                 140

Cys Lys Asp Lys His Arg Phe Asp Gln Lys Asn Ser Asn Thr Tyr Val
145                 150                 155                 160

Lys Thr Asn Lys Thr Trp Ala Ile Ala Tyr Gly Thr Gly Asp Ala Arg
                165                 170                 175

Gly Phe Phe Gly Arg Asp Thr Val Arg Leu Gly Ala Glu Gly Lys Asp
            180                 185                 190

Gln Leu Val Ile Asn Asp Thr Trp Phe Gly Gln Ala Glu His Ile Ala
        195                 200                 205

Glu Phe Phe Ser Asn Thr Phe Leu Asp Gly Ile Leu Gly Leu Ala Phe
    210                 215                 220

Gln Glu Leu Ser Glu Gly Gly Val Ala Pro Pro Ile Ile Arg Ala Ile
225                 230                 235                 240

Asp Leu Gly Leu Leu Asp Gln Pro Ile Phe Thr Val Tyr Phe Glu Asn
                245                 250                 255

Val Gly Asp Lys Glu Gly Val Tyr Gly Gly Val Phe Thr Trp Gly Gly
            260                 265                 270

Leu Asp Pro Asp His Cys Glu Asp Glu Val Thr Tyr Glu Gln Leu Thr
```

```
        275                 280                 285

Glu Ala Thr Tyr Trp Gln Phe Arg Leu Lys Gly Val Ser Ser Lys Asn
    290                 295                 300

Phe Ser Ser Thr Ala Gly Trp Glu Ala Ile Ser Asp Thr Gly Thr Ser
305                 310                 315                 320

Leu Asn Gly Ala Pro Arg Gly Ile Leu Arg Ser Ile Ala Arg Gln Tyr
                325                 330                 335

Asn Gly Gln Tyr Val Ala Ser Gln Gly Leu Tyr Val Val Asp Cys Ser
            340                 345                 350

Lys Asn Val Thr Val Asp Val Thr Ile Gly Asp Arg Asn Tyr Thr Met
        355                 360                 365

Thr Ala Lys Asn Leu Val Leu Glu Ile Gln Ala Asp Ile Cys Ile Met
    370                 375                 380

Ala Phe Phe Glu Met Asp Met Phe Ile Gly Pro Ala Trp Ile Leu Gly
385                 390                 395                 400

Asp Pro Phe Ile Arg Glu Tyr Cys Asn Ile His Asp Ile Glu Lys Lys
                405                 410                 415

Arg Ile Gly Phe Ala Ala Val Lys His
            420                 425
```

<210> 7
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Synthetic peptide

<400> 7

```
Ala Leu Glu Arg Thr Phe Leu Ser Phe Pro Thr
  1               5                  10
```

INTERNATIONAL SEARCH REPORT

Inter tional Application No
PCT/GB 01/00819

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K16/40 C12N9/64 A61K39/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K A61K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EMBL, BIOSIS, WPI Data, EPO-Internal, PAJ, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BROWN ALAN ET AL: "Necator americanus (human hookworm) aspartyl proteinases and digestion of skin macromolecules during skin penetration." AMERICAN JOURNAL OF TROPICAL MEDICINE AND HYGIENE, vol. 60, no. 5, May 1999 (1999-05), pages 840-847, XP001000919 ISSN: 0002-9637 abstract page 846, column 1, paragraph 3 -column 2, paragraph 2 --- -/--	12-14

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 July 2001	31/07/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax. (+31-70) 340-3016	Montrone, M

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	DAUB J ET AL: "A survey of genes expressed in adults of the human hookworm, Necator americanus." PARASITOLOGY, vol. 120, no. 2, February 2000 (2000-02), pages 171-184, XP001000928 ISSN: 0031-1820 abstract page 171, column 1, line 10-12 -column 2, line 15-18 page 173; table 1 page 179, column 2, paragraph 2 ---	12-14
Y	BROWN A ET AL: "An initial characterization of the proteolytic enzymes secreted by the adult stage of the human hookworm Necator americanus." PARASITOLOGY, vol. 110, no. 5, 1995, pages 555-563, XP001000918 ISSN: 0031-1820 abstract page 555, column 1, paragraph 3 -column 2, paragraph 1 page 560, column 1, paragraph 2 -column 2, paragraph 6 page 561, column 2, paragraph 2 ---	12-14
Y	DATABASE EMBL 'Online! Accession No.: AI857115, 22 July 1999 (1999-07-22) DAUB J.: XP002172116 abstract ---	12-14
P,X	DATABASE EMBL 'Online! Accession No.: Q9N9H4, 1 October 2000 (2000-10-01) GIRWOOD ET AL: XP002172117 abstract ---	13,14
P,X	DATABASE EMBL 'Online! Accession No.: NAM245458, 27 July 2000 (2000-07-27) GIRWOOD ET AL: XP002172118 abstract ---	13,14
P,X	DATABASE EMBL 'Online! Accession No.: AJ245459, 27 July 2000 (2000-07-27) GIRWOOD ET AL: XP002172119 abstract -----	13,14

1